



Getting Started with the ECLS-K Data

Module Objectives

- Introduce users to the data collected in the ECLS-K and the resulting micro-level data files
- Describe the resources available to learn more about the study, the data, and the data files

ECLS-K Data

ECLS-K data are obtained directly from respondents across multiple points in time in the study

Respondents	Data Collection Rounds	Collection Methods
<ul style="list-style-type: none">• Children• Parents or guardians• Teachers• School administrators	<ul style="list-style-type: none">• Fall and spring of kindergarten• Fall and spring of first grade• Spring of third grade• Spring of fifth grade• Spring of eighth grade	<ul style="list-style-type: none">• Direct assessment• Telephone interviews• Self-administered, hard-copy questionnaires

Note: [Data collection instruments](#)

Other Categories of Available ECLS-K Data

- Administrative variables
 - Information about study respondents and different aspects of data collection compiled and maintained in administrative records throughout the course of the study
- Composite variables
 - Variables created using information from two or more variables, two or more sources, or both
 - Sometimes derived using information that is not available to researchers in the data files
- Weighting variables
 - Variables used in the process of weighting and standard error adjustment
- Access the [Data File User's Manuals](#)

Data File Basics

- Data are available in both public- and restricted-use formats
 - Most of the available public-use ECLS-K data are included in the Kindergarten Through Eighth Grade Full Sample Public-Use Data File. Note that there is an error in one set of scores on this file; users should review the [errata](#) describing this error
 - Restricted-use data are available in round-specific cross-sectional files and supplementary data files
 - Restricted-use file for a given round of data collection contains more data and a wider range of data values than are included for that round in the public-use files
 - All values for all cases for restricted variables are set to -2 in the public-use file
 - Data file user's manual table in chapter 7 lists all the variables that have been altered in some way
- Both public- and restricted-use data come on a CD or DVD containing the data and data documentation

Data File Basics (Continued)

- The CD/DVD also contains Electronic Codebook software, or an ECB, which is software that helps users to create customized data files with only the variables needed for analysis
 - View descriptions of variables, variable categories, and frequency distributions
 - Identify variables to include in a customized data file
 - Create a syntax file to be run in SAS, SPSS, or Stata to generate a data file
 - The ECB does not produce a data file
- Do **not** alter the ASCII data files, which end with a .dat extension
- Public-use data can also be downloaded as full data files from the [ECLS-K website](#). An online data tool called the [EDAT](#), can also be used to help researchers create customized data files containing public-use data.

ECLS-K Data Files Available for Analysis, Public-use Files

- Kindergarten through eighth-grade full sample public-use data file
- Base-year teacher-level file
- Base-year school-level file
- Supplementary data files containing item-level data for the Self-Description Questionnaire (SDQ) and Approaches to Learning items

ECLS-K Data Files Available for Analysis, Restricted-use Files

- Cross-sectional restricted-use data files (kindergarten, first, third, fifth, and eighth grades)
- Head Start verification data (base year)
- Kindergarten teacher and school administrator salary and benefits data (base year)
- Special education teacher data (base year)
- Student record abstract (base year)
- Geocode data file (kindergarten through third grade)
- Parent and teacher Social Rating Scale (SRS) items and data files (kindergarten through fifth grade)

Variable Naming Conventions

First character generally indicates the data collection instrument
(e.g., **P** = Parent interview, **S** = School administrator questionnaire)

Second character generally indicates the round of data collection
(i.e., **1** = Fall kindergarten, **2** = Spring kindergarten, **3** = Fall first grade,
4 = Spring first grade, **5** = Spring third grade, **6** = Spring fifth grade,
7 = Spring eighth grade)

Exception: The second character of composite variables beginning
with **W** indicates grade, not round.
(i.e., **K** = Kindergarten, **1** = First grade, **3** = Third grade, **5** = Fifth grade,
8 = Eighth grade)

Variable Naming Conventions, Examples

P2MEETTM:

P = Parent interview

2 = Round 2 (spring kindergarten) data collection

S6GFTPRG:

S = School administrator questionnaire

6 = Round 6 (spring fifth grade) data collection

Click here to access the [Variable Naming Conventions](#) resource sheet.

Variable Label Conventions

- First character generally indicates the collection instrument (e.g., P = Parent interview, S = School administrator questionnaire)
- Second character generally indicates the round of data collection (i.e., 1 = Fall kindergarten, 2 = Spring kindergarten, 3 = Fall first grade, 4 = Spring first grade, 5 = Spring third grade, 6 = Spring fifth grade, 7 = Spring eighth grade)
 - Exception: The second character of composite variables beginning with W indicates grade, not round (i.e., K = Kindergarten, 1 = First grade, 3 = Third grade, 5 = Fifth grade, 8 = Eighth grade)
- Next block of characters describes the variable
- Labels
 - Labels for variables associated with specific questions asked directly of parent interview respondents, teachers, and school administrators include the question number
 - Labels for parent interview variables also have identifiers for the section of the interview in which the question was asked
 - The labels also include some text related to the actual question wording

Variable Label Conventions (Continued)

Example: P4EVALUA: P4 CHQ040 CHD LEARNING ABILITY EVALUATED

P = Parent interview

4 = Round 4 (spring first-grade) data collection

CHQ = Child Health section of the parent interview

040 = The question number

CHD LEARNING ABILITY EVALUATED= Paraphrasing of the actual question wording

Variable Label Conventions (Continued)

Descriptions used in labels for variables that are not associated with specific questions asked directly of respondents are varied in their format

Example: **F2SPECS: F2 CHILD RECEIVED SPEC EDU SERV FROM FMS**

F = Variable derived from administrative data in the field management system (FMS)

2 = Round 2 (spring kindergarten) data collection

F2 CHILD RECEIVED SPEC EDU SERV FROM FMS = Description of what the variable indicates

ECLS-K Data File Organization

The K-8 full sample public-use data file is organized into blocks of data corresponding to specific content areas

- Identification numbers (child, parent, school, teacher)
- Administrative data from the Field Management System (F2, F4, F5, F6, and F7 variables)
- School characteristics data derived from NCES school universe data files
- Basic child information (e.g., sex, race, age at time of assessment)
- Full sample weights
- Child assessment scores
- Composite variables
- Child assessment item-level data and child questionnaire data (C1, C2, C3, C4, C5, C6, and C7 variables)
- Parent interview data (P1, P2, P3, P4, P5, P6, and P7 variables)
- Classroom teacher – teacher and classroom data (kindergarten through grade three) (A1, A2, A4, A5, B1, B2, B4, and B5 variables)
- Classroom teacher – child-specific data (kindergarten through grade three) (T1, T2, T4, and T5 variables)
- Classroom teacher – teacher and classroom data (grades five and eight) (J61, J62, J71, and J72 variables)
- Classroom teacher – child-specific data (grades five and eight) (G6, G7, M6, M7, N6, and N7 variables)
- School fact sheet data (L5 variables)
- School administrator questionnaire data (S2, S4, S5, S6, and S7 variables)
- School facilities checklist data (K2, K4, K5, and K6 variables)
- Special education teacher, teacher data (D2, D4, D5, D6, and D7 variables)
- Special education teacher, child-specific data (E2, E4, E5, E6, and E7 variables)
- Student record abstract data (U2, U4, U5, and U6 variables)
- Replicate weights
- Taylor Series sampling strata and primary sampling unit variables

Resources for Information on ECLS-K Data

- Read the [user's manuals and other technical reports](#) such as psychometric reports
 - Data considerations, caveats, anomalies, errata
- Review substantive reports such as initial release reports
- Review the [study bibliography](#)
- Review instruments
 - Check skip patterns
- Check the website occasionally

Module Summary and Resources

- Described the data collected through the ECLS-K and the resulting micro-level data files
- Described the resources available to learn more about the study, the data, and the data files

Resources

- [ECLS-K Data Collection Instruments](#)
- [Data File User's Manuals and other technical reports](#)
- [Errata](#)
- [ECLS-K Data Products](#)
- [EDAT](#)
- [Variable Naming Conventions resource sheet](#)
- [Study Bibliography](#)